1 **METHOD AND SYSTEM FOR IDENTIFYING MANUFACTURING ANOMALIES IN A**

2 **MANUFACTURING SYSTEM**

3

4 **BACKGROUND OF THE INVENTION**

5 1.   Field of the Invention

6   The invention relates to a method for identifying manufacturing anomalies in a

7 manufacturing system comprising a plurality of products which are manufactured with a plurality

8 of manufacturing attributes. More specifically, the invention comprises a method and system for

9 using data-mining techniques to define normality for product performance, and to identify

10 anomalies for manufacturing parameters excluded from the defined normality based on product

11 performance.

12 2.   Description of the Prior Art and Related Information

13   Today's ever smaller manufactured products may comprise multitudes of components

14 which are assembled in manufacturers' manufacturing facilities. In order to secure steady

15 supplies and obtain lower pricing on components, manufacturers of the products may use a

16 plurality of sources for the components. Such a diversification technique may include receiving

17 different groups, or lots, of the same component from different manufacturers to be included in

18 the same product line or model.

19   Typically, components are shipped from the component suppliers in lots that have similar

20 component manufacturing parameters. Such parameters may comprise, for example, direct

21 parameters such as the tolerance of a lot of resistors, the threshold of a lot of resistors, the

22 capacitance of a lot of capacitors, the reactance of a log of capacitors; or indirect parameters such

23 as the specific supplier from which a particular lot originated; the shipping method used for

24 transporting the lot to the product manufacturer, or the time of year (date code) that the lot of

25 components were manufactured. Such direct parameters may cause a variation in the quality,

26 reliability or performance of a particular lot of components, thereby causing a variation in the

27 performance of the assembled product. Further, correlation may exist between a variation in

1   performance and indirect parameters. Further, there may be certain interactions between

2   parameters that may lead to variations in performance. Such variations may be positive or

3   negative influences on product performance.

4   The lots are typically identified by a lot number, called a component identifier herein,

5   which identifies the lot to the product manufacturer and the supplier, the component identifier

6   usually being used for accounting purposes or traceability to manufacturing parameters of the lot.

7   However, especially with small electronic components, the component identifiers themselves

8   may not be imprinted or bar-coded on the components themselves due to practical considerations

9   such as component size, or the extra time and expense in manufacturing that would be required

10 to do so. Therefore, once the components leave their packing materials in which they were

11 shipped, the lot from which those components came from may not be able to be identified.

12   Certain lots of components may be in some way defective or vary in their performance

13 due to the different manufacturing conditions referenced above, or due to a certain component

14 supplier's neglect or breach in promise to supply a certain quality of components. One solution

15 that has been employed heretofore is for the manufacturer to test statistically significant numbers

16 of components from each lot as they arrive from the individual suppliers. However, with

17 products that are sold for low profit margins, such testing and delay in product assembly has

18 become untenable. More and more product manufacturers have been forced to rely on

19 component suppliers' representations that the supplied components meet the standards set by the

20 product manufactures, or that the supplied components fall within a set standard deviation of

21 performance within each lot. Such reliance has proven inadequate in countless situations, with

22 many product manufactures having their products show variations in performance when it is too

23 late to trace the suspect components back to the lot from which they came.

24   Still, apart from the components themselves, there may be product manufacturing

25 parameters which may be responsible for variations in performance in products. For example,

26 the temperature or humidity of a factory where the components are assembled into the products

1   may be a contributing factor to product performance. Other product manufacturing parameters

2   such as the assembly line or conveyor belt speed my further affect product performance.

3   Solutions to the above problems are described in U.S. Patent Application Serial No.

4   09/494,175, filed January 31, 2000, entitled METHOD USING STATISTICALLY ANALYZED

5   PRODUCT TEST DATA TO CONTROL COMPONENT MANUFACTURING PROCESS

6   which is hereby incorporated by reference herein. That application provides for a plurality of

7   analytical tools for identifying correlations between products which show variations in

8   performance and product or component manufacturing parameters. Typical analytical tools

9   which are described in the system of that application include statistical data tools, a data

10   visualization tool, a data-mining tool, an on-line analytical processing (OLAP) tool, an

11   information broadcast tool or some combination thereof. However, typical analytical tools

12   require that the operator of such tools have a working knowledge of which manufacturing

13   parameters may cause product performance to suffer. The programmer is then required to

14   program the operation of a particular analytical tool based on predicted manufacturer parameter

15   anomalies. As a result, unpredictable anomalies may thus go undetected, or may be difficult to

16   pinpoint.

17   Finally, it is often difficult to ascertain normal operation of a product relative to a

18   manufacturing parameter as compared with anomalous operation. This is especially so with

19   respect to multi-variant manufacturing parameters which may have anomalies in one or more of

20   several performance measure dimensions. For example, one or more configurations of a disk

21   drive head assembly may have anomalous operation with respect to seek time while having a

22   relatively low error rate and normal servo tracking. Given that other head assembly

23   configurations may show anomalous performance in seek time, error rate or servo tracking,

24   defining just what is normal operation can be difficult, much less identifying anomalous

25   operation.

26   Thus, there is a need for a system and method for defining normal operation in product

27   performance with respect to manufacturing parameters in multiple dimensions. There is further a

1   need for a system and method for identifying unpredicted or unpredictable manufacturing

2   anomalies.

3   **SUMMARY OF THE INVENTION**

4   A method and system for identifying manufacturing anomalies in a manufacturing system

5   comprising a plurality of products which are manufactured with a plurality of manufacturing

6   parameters is disclosed. The system comprises a data mining program applied to the data

7   warehouse for analyzing the stored manufacturing parameters to define a normal manufacturing

8   parameter subset. The data mining program is further for detecting that at least one of the

9   plurality of manufacturing parameters is excluded from the normal subset. The system further

10 comprises a reporting means for reporting the at least one detected manufacturing parameter.

11   The data mining program may further detect that a plurality of the manufacturing

12 parameters are excluded from the first subset. The data mining program then analyzes the

13 detected plurality of manufacturing parameters to define a second normal subset of the detected

14 plurality of manufacturing parameters. The reporting means may then report the second normal

15 subset of manufacturing parameters.

16   The data mining program defines a normal subset by first classifying performance

17 variables in n-dimensional space. The manufacturing parameters are plotted against various

18 performance values. The data mining program defines one or more performance clusters, or

19 subsets, of parameters of products which show similar performance ratings in the n-dimensional

20 space. The data mining program is then able to detect parameters which fall outside the groups,

21 thereby defining an anomaly in performance for that parameter.

22 **BRIEF DESCRIPTION OF THE DRAWINGS**

23 Fig. 1 is a flow diagram illustrating the steps performed by the system of the present invention;

24 Fig. 2 is a block diagram showing an exemplary structure of a system that performs the method

25 illustrated in Fig. 1;

26 Fig. 3 is a simplified n-dimensional space plot of manufacturing parameters based on product

27 performance produced by the data mining program of the system of Fig. 2;

1   Fig. 4 is a simplified n-dimensional space plot alternative to that shown in Fig. 3;

2   Figs. 5-6 are visual representations of one of the methodologies used in the overall categorization

3   performed by the data mining program of Fig. 2; and

4   Fig. 7 is an output from another visualization tool of the data mining program 54 similar to that

5   illustrated in Fig. 4.

6   **DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS**

7   With reference to Fig. 1, a flow diagram illustrating the steps performed in a method for

8   identifying manufacturing anomalies in a plurality of products which are manufactured with a

9   plurality of manufacturing parameters in a manufacturing system described in Fig. 2 referenced

10   below is shown. The plurality of manufacturing parameters (20 in Fig. 2) are stored in the data

11   warehouse (56 in Fig. 2) step 300. A data mining program (54 in Fig. 2) is applied to analyze

12   the stored manufacturing parameters 20 to define a first normal manufacturing parameter subset

13   (102 in Fig. 3 described below), step 302. The first normal subset 102 is defined by categorizing

14   the manufacturing parameters 20 in n-dimensional space (100 in Fig. 3). The data mining

15   program 54 detects at least one of the plurality of manufacturing parameters 20 that is excluded

16   from the first normal subset 102 in the n-dimensional space 100, step 304. The detection of the

17   at least one manufacturing parameter 20 is then reported to personnel through a notification

18   system (70 in Fig. 2), step 306.

19   The data mining program 54 may detect that a plurality of the manufacturing parameters

20   20 are excluded from the first normal subset 102, step 308. In this case, the data mining program

21   54 may analyze the detected plurality of manufacturing parameters 20 to define a second normal

22   subset (106 in Fig. 3) of the detected plurality of manufacturing parameters 20, step 110. The

23   second normal subset 106 is defined by the data mining program 54 by categorizing the

24   manufacturing parameters 20 excluded from the first normal subset 102 in the n-dimensional

25   space 100. The notification system 70 then reports the second normal subset 106 of

26   manufacturing parameters 20 to personnel, step 312.

1   With reference to Fig. 2, an exemplary structure of the system for identifying

2   manufacturing anomalies in a manufacturing system is shown. The system includes a

3   manufacturing information system 51 for a manufacturer 50 such as the above-referenced U.S.

4   Patent Application entitled METHOD USING STATISTICALLY ANALYZED PRODUCT

5   TEST DATA TO CONTROL COMPONENT MANUFACTURING PROCESS. The system

6   includes a data warehouse 57 for storing product identifiers 56 identifying products having

7   manufacturing parameters 20 and values representing the manufacturing parameters 20. The

8   manufacturing parameters 20 may comprise product manufacturing parameters and/or

9   component manufacturing parameters.

10   A data mining program 54 is applied to the data warehouse for analyzing the stored

11   manufacturing parameters 20 to define the first normal subset 102 of manufacturing parameters

12   20. The data mining program 54 detects that at least one of the manufacturing parameters 20 is

13   excluded from the normal subset 102.

14   A reporting means, comprising a notification system 70, is included for reporting the at

15   least one detected manufacturing parameter 20 to personnel.

16   A relational means 58 is provided in the data warehouse for relating the manufacturing

17   parameters 20 with particular product identifiers 56 for products that have the related

18   manufacturing parameters 20. The relational means 58 may comprise an RDBMS system such

19   as that produced by ORACLE Systems of Redwood Shores, California or by NCR of Dayton,

20   Ohio.

21   The system includes a means for receiving 60 product performance data 80, wherein the

22   performance data 80 includes product identifier data for matching with the product identifiers 56

23   in the data warehouse 57. The means for receiving 60 may comprise a data terminal and

24   keyboard, personal computer or failure analysis test equipment for receiving data from failed

25   products.

26   The data mining program 54 is applied to the data warehouse 57 for defining a normal

27   subset by first classifying performance variables in n-dimensional space as described with

1 respect to Fig. 3 below. The manufacturing parameters 20 are plotted against various

2 performance values. The data mining program 54 defines one or more performance clusters, or

3 subsets (102-106 in Fig. 3), of manufacturing parameters 20 of products which show similar

4 performance ratings in the n-dimensional space. The data mining program 54 is then able to

5 detect manufacturing parameters 20 which fall outside, or are excluded from, a subset 102,

6 thereby defining an anomaly in performance for that manufacturing parameter 20. Further, a

7 new subset 106 may be defined by the data mining program 54 from performance clusters

8 defined by a plurality of manufacturing parameters 20 excluded from the first normal subset 102.

9 The data mining computer program 54 may comprise a plurality of data mining sub-

10 routines 59 for executing. One or more of the sub-routines 59 may be executed to define one or

11 more first normal subsets 102 of the manufacturing parameters 20. Another data mining sub-

12 routine 59 may be used to detect that one or more parameters 20 fall outside the one or more

13 normal subsets. Another sub-routine 59 may be executed for analyzing the detected

14 manufacturing parameters 20 to define new normal subsets 106 from the detected plurality of

15 manufacturing parameters 20. Each sub-routine 59 may be divided between a plurality of

16 executable files for executing against the data warehouse 57.

17 With reference to Fig. 3, a simplified plot in n-dimensional space 100 of manufacturing

18 parameters 20 based on product performance produced by the data mining program 54 of the

19 method of Fig. 1 and system of Fig. 2 is shown. For the n-dimensional space 100 of Fig. 3, n=3

20 in order to simplify the diagram for illustrative purposes. The dimensions may represent error

21 rate, access speed, and yield for disk head assemblies. Each axis in the n-dimensional space 100

22 of Fig. 3 represents the scale for a dimension. As described above, the data mining program 54

23 has defined a normal subset 102 based on performance data 80 received from the input means 60

24 plotted against each axis. At locations 104 in the n-dimensional space 200, a plurality of

25 performance parameters 20 are separated from the normal subset 102. The data mining program

26 54 has further defined other normal subsets 106 in the n-dimensional space 100 with

27 performance parameters that are separated from the normal subset 102.

1   With reference back to Fig. 2, a notification system 70 causes persons to be notified when

2   either a performance parameter 20 (location 104 in Fig. 3) is excluded from the normal subset

3   102, or a new normal subset 106 (locations 106 in Fig. 3) is found by the data mining program

4   54. When the system requires reporting to users, a personnel notification identification database

5   (PNID) 72 is consulted. In the embodiment of Fig. 2, the PNID 72 comprises a plurality of

6   PNID parameter records 20a that each match up to one of the parameter records 20 in the data

7   warehouse 57. If a manufacturing parameter 20 is identified as being excluded from the normal

8   subset 102, or a second normal subset 106 is identified and defined by the data mining program

9   54, then the system matches all of the identified manufacturing parameters 20 with the

10   corresponding PNID parameter records 20a in the PNID 72.

11   The PNID 72 further contains personnel records 76. The personnel records 76 each

12   identify one or more persons to be notified when one or more PNID parameter records 20a are

13   matched with the identified manufacturing parameters 20. The PNID parameter records 20a are

14   related to the personnel records 76 through a relational means 78 such as that described above

15   with respect to relational means 58.

16   A notification method database (NMD) 82 within the notification system 70 is next

17   consulted to determine the method 700 by which each person identified by each personnel record

18   76 should be notified. The NMD 82 contains notification method personnel records (NMPRs)

19   76a stored therein which are readily matched with the personnel records 76 of the PNID 72.

20   Each NMPR 76a is related with at least one of a plurality of method records 702a-708a through a

21   relational means 88 such as that described above with respect to relational means 58 above. Each

22   method record 702a-708a designates a method of notification 700 which the notification system

23   72 acts upon to provide notification to personnel. For each person who is to be notified, as

24   indicated by matching the personnel records 76 that are related to PNID parameter records 20a

25   that are matched with identified manufacturing records 20, the related method records 702a-708a

26   indicate to the notification system 70 how each person should be notified. In the embodiment of

27   Fig. 2, method record 702a indicates that electronic mail 702 should be used to notify the

1   specified person, method record 704a indicates that voice mail 704 should be provided, method

2   record 706a indicates that a page 706 should be issued and method record 708a indicates that a

3   person should be notified by cell or wireless phone 708.

4   The above described actions of the notification system 82 may be executed by a

5   combination of hardware in conjunction with a notification program 84. A personal computer

6   with an RDBMS stored therein may comprise the notification system 70.

7   With reference to Fig. 4, another embodiment of a simplified plot of manufacturing

8   parameters 20 in n-dimensional space 100 based on product performance produced by the data

9   mining program 54 of the method of Fig. 1 and system of Fig. 2 is shown. For the n-dimensional

10   space 100 of Fig. 3, n=3 in order to simplify the diagram for illustrative purposes. As with the n-

11   dimensional space 100 of Fig. 3, the dimensions may represent error rate, access speed, and yield

12   for disk head assemblies. Each axis in the n-dimensional space 100 of Fig. 3 represents the scale

13   for a dimension. The data mining program 54 has defined a normal subset 102 based on

14   performance data 80 received from the input means 60 plotted against each axis. However, in

15   Fig. 3, the normal subset 102 forms a surface instead of a cluster. At locations 104 in the n-

16   dimensional space 100, a plurality of performance parameters 20 are separated from the normal

17   subset 102 surface. The data mining program 54 has further defined other normal subset 106

18   surfaces in the n-dimensional space 100 with performance parameters that are separated from the

19   normal subset 102 surface.

20   Fig. 5 is a visual representation of one of the methodologies used in the overall

21   categorization performed by the data mining program 54 to determine normal and abnormal

22   subsets. This particular form of the data mining program 54 presents a categorization that is

23   called a decision tree because of branching that takes place to categorize and subset groups of the

24   manufacturing parameters 20. A base root node of the tree (barplot at the front in Fig. 5) shows

25   the response variable (e.g., yield) plotted by quantity within bins (e.g, number of products with

26   yields from 0-5%, 5-10%, ...95-100%). At the base of this root node, the data mining program

27   54 lists a most significant manufacturing parameter (in the n dimensional space 100) for

1   grouping the yield into distinct (non overlapping) groups (similar to groups 102 and 106 in figure

2   4). The conditions applied for the grouping are published by the data mining program 54 (e.g.,

3   total product returns <= 18 for the left branch, total returns >18 for the right branch) - and are

4   shown in Fig. 6 for one of the branches. Fig. 6 also shows the distribution of a subset (note this

5   subset consists strictly of high valued devices – no low valued ones are present).

6   This process is repeated for each subset until the data mining program determines 54 that

7   no statistically significant improvement in segregation can be made. Thus in terms of Fig. 4,

8   there would essentially be 7 terminal nodes representing groups 102, 106 and up to 5 individual

9   nodes representing the independent outliers represented by 104 in figure 4.

10   Fig. 7 is an output from another visualization tool of the data mining program 54 similar

11   to that illustrated in Fig. 4. The plot in Fig. 7 illustrates up to 5 dimensions by varying the x,y,z

12   positions, the color of the clusters, or balls, and the opacity of the balls. In addition the size of

13   the balls is proportional to the quantity of units in each subset 102, 106. This visualization tool

14   may be used by the investigator to verify the statistical groupings made by the data mining

15   program 54 like in Fig. 4.